

On a Construction of Entropic Vectors Using Lattice-Generated Distributions

Babak Hassibi and Sormeh Shadbakht

EE Department

California Institute of Technology

Pasadena, CA 91125

hassibi,sormeh@caltech.edu

Abstract—The problem of determining the region of entropic vectors is a central one in information theory. Recently, there has been a great deal of interest in the development of non-Shannon information inequalities, which provide outer bounds to the aforementioned region; however, there has been less recent work on developing inner bounds. This paper develops an inner bound that applies to any number of random variables and which is tight for 2 and 3 random variables (the only cases where the entropy region is known). The construction is based on probability distributions generated by a lattice. The region is shown to be a polytope generated by a set of linear inequalities. Study of the region for 4 and more random variables is currently under investigation.

I. INTRODUCTION

Let X_1, \dots, X_n be a collection of n jointly distributed finite-alphabet random variables and consider the $2^n - 1$ dimensional vector whose entries are the joint entropies of each non-empty subset of these n random variables. Any $2^n - 1$ dimensional vector that can be constructed from the entropies of n such random variables is referred to as *entropic*. Characterizing the region of entropic vectors has long been an interesting open problem. Many issues in information theory and probabilistic reasoning such as optimizing information quantities or characterizing the compatibility of conditional independence relations involve, or are closely related, to this problem. Moreover, characterizing this region is fundamental in the sense that many network information theory problems can be formulated as linear optimization problems over this region. Thus, determining this region can lead to the solution of a whole host of information-theoretic problems. On the other hand many proofs of the converse of coding theorems involve information inequalities, the complete set of which can be found as a result of characterizing this region.

In [1], Yeung refers to the space of entropic vectors as Γ_n^* . The work of Han, Fujishige, Zhang and Yeung, [2], [3], [4], [5], [1] has resulted in the complete characterization of Γ_n^* for $n = 2, 3$ and their relation to polymatroids and submodular functions. In particular, if we let $\mathcal{N} = \{1, \dots, n\}$, $\alpha, \beta \subseteq \mathcal{N}$, $X_\alpha = \{X_i : i \in \alpha\}$ and $X_\beta = \{X_i : i \in \beta\}$, it is clear that the entropy $H(X_\alpha) = H(\alpha)$ (for simplicity) satisfies the properties

- 1) $H(\emptyset) = 0$
- 2) For $\alpha \subseteq \beta$: $H(\alpha) \leq H(\beta)$
- 3) For any α, β : $H(\alpha \cup \beta) + H(\alpha \cap \beta) \leq H(\alpha) + H(\beta)$

the last of which is referred to as the submodularity property. The above inequalities are referred to as the basic inequalities of Shannon information measures (and are derived from the positivity of conditional mutual information). Any inequalities that are obtained as positive linear combinations of these are simply referred to as *Shannon inequalities*. The space of all vectors of $2^n - 1$ dimensions whose components satisfy all such Shannon inequalities is denoted by Γ_n . It has been shown that [5] $\Gamma_2^* = \Gamma_2$ and $\bar{\Gamma}_3^* = \Gamma_3$, where $\bar{\Gamma}_3^*$ is the closure of Γ_3^* . However, for $n = 4$, recently several “non-Shannon-type” information inequalities have been found that demonstrate that Γ_4^* is *strictly* smaller than Γ_4 [5], [6], [7], [8], [9].

Although these non-Shannon type inequalities give outer bounds for Γ_4^* , its complete characterization remains an open problem. The non-Shannon inequalities have also proven useful in deriving various outerbounds for different network information theory problems [10], [11]. However, there has been much less focus on determining *inner bounds* on Γ_n^* [4], [12]. These would be of great interest since they would yield achievable rate regions for many network information theory problems. The focus of this paper will therefore be on determining an inner bound on Γ_n^* with two goals in mind: an inner bound that can be extended to *any* n , and that can be easily expressed (the inner bound that we shall later give is a polytope, i.e., defined by a set of linear inequalities).

We should mention that in this paper, we shall focus on *normalized* entropy vectors which we define as follows:

$$h(\alpha) = \frac{1}{\log N} H(X_\alpha), \quad (1)$$

where N is the alphabet size of the X_i (it can be assumed wlog that the alphabet size is equal for all the X_i by simply adding letters with zero probability wherever necessary).

Definition 1 (Normalized-Entropic Vectors): Any vector $v \in \mathcal{R}^{2^n - 1}$ which can be regarded as the normalized entropy vector of some collection of n random variables, for some value of N , is called normalized-entropic (or just entropic when there is no ambiguity). We will denote the space of normalized-entropic vectors by Ω_n^* .

There are several reasons for considering this normalized version: it is often the normalized version that comes up in capacity calculations (where the normalization represents the number of channel uses), it makes the entropy region finite,

and makes the proof of the convexity of the closure of Ω_n^* trivial. Indeed if $h(X_\alpha)$ and $h(Y_\beta)$ are the entropies of the X_i and Y_i , each with alphabet size N_x and N_y , then by making n_x independent copies of the X 's and n_y independent copies of the Y 's, so that the overall alphabet size becomes $N_x^{n_x} N_y^{n_y}$, we obtain the normalized entropy

$$\frac{n_x \log N_x}{n_x \log N_x + n_y \log n_y} h(X_\alpha) + \frac{n_y \log N_y}{n_x \log N_x + n_y \log n_y} h(Y_\beta), \quad (2)$$

which since n_x and n_y are arbitrary, implies that one can get arbitrary close to any point on the convex hull of h_x and h_y and this implies the convexity of the closure of Ω_n^* .

The following theorem states the connection between Ω_n^* and Γ_n^* , proof of which we omit for brevity:

Theorem 1 (Ω_n^ and Γ_n^*):* Define the ray of a set S as

$$\text{ray}(S) = \{\alpha X | \alpha > 0, X \in S\}. \quad (3)$$

Then we have

$$\text{ray}(\bar{\Omega}_n^*) = \bar{\Gamma}_n^*, \quad (4)$$

i.e., the ray of $\bar{\Omega}_n^*$ is the closure of Γ_n^* .

The remainder of the paper is organized as follows. The next section studies quasi-uniform distributions, which will be the building blocks for our construction. Section III contains the main results of our paper, especially the construction of entropic vectors using lattice-generated probability distributions. Section IV makes the construction explicit for 2, 3 and 4 random variables and shows that it achieves the full entropy region for $n = 2, 3$.

II. QUASI-UNIFORM DISTRIBUTIONS

One way of characterizing Γ_n^* is through determining its kissing hyperplanes:

$$a^T H = \sum_{\alpha \subseteq \mathcal{N}} a_\alpha H_\alpha \geq \gamma, \quad (5)$$

for $a \in \mathcal{R}^{2^n-1}$ and for all $H \in \Gamma_n^*$. To determine the value of γ , one needs to perform the optimization

$$\gamma = \min_{H \in \Gamma_n^*} \sum_{\alpha \subseteq \mathcal{N}} a_\alpha H_\alpha. \quad (6)$$

One of the difficulties in performing this optimization is that the alphabet size of the underlying distribution is arbitrary. Nonetheless, if we restrict the alphabet size of each X_i to N and attempt to optimize over the unknown joint distribution $p_{X_{\mathcal{N}}}(x_{\mathcal{N}})$ then the KKT conditions necessitate that

$$\sum_{\alpha \subseteq \mathcal{N}} a_\alpha \log \frac{1}{p_{X_\alpha}(x_\alpha)} = c \quad \text{if } p_{X_{\mathcal{N}}}(x_{\mathcal{N}}) \neq 0, \quad (7)$$

for some constant c . The KKT conditions imply that, rather than searching over all possible distributions $p_{X_{\mathcal{N}}}(x_{\mathcal{N}})$, we need only search over those distributions that satisfy (7).

Of course, there can be many solutions to (7). However, a rather obvious solution—and one that does not depend on a ,

the normal vector of the hyperplane—is the following: for any $\alpha \subseteq \mathcal{N}$:

$$p_{X_\alpha}(x_\alpha) = c_\alpha \quad \text{or } 0 \quad (8)$$

for some constant c_α , independent of the point $x_\alpha \in \{1, \dots, N\}^{|\alpha|}$. In other words, these are distributions that take on zero or a constant value for all possible marginals, $p_{X_\alpha}(\cdot)$. Such distributions are referred to as *quasi-uniform* [13].

Computing the entropy for quasi-uniform distributions is, of course, straightforward:

$$H(\alpha) = \log \frac{1}{c_\alpha}. \quad (9)$$

Let Λ_n denote the space of entropy vectors generated by quasi-uniform distributions. Then the remarkable result of [14], [13] is that

Theorem 2 (Quasi-Uniform Distribution): $\overline{\text{con}}(\Lambda_n) = \bar{\Gamma}_n^*$, i.e., the convex closure of Λ_n is the closure of Γ_n^* .

In other words, considering quasi-uniform distributions is sufficient for characterizing $\bar{\Gamma}_n^*$. As a result, these are the distributions we will henceforth focus on.

III. DISTRIBUTIONS FROM LATTICES

Determining all quasi-uniform distributions appears to be a hopelessly complicated combinatorial problem. Since we are looking for a construction that can be generalized to any n , it seems reasonable that we need to impose some structure. Some circumspection suggests the use of a lattice structure.

Recall that we can generally represent the points generated by a lattice in an n dimensional space as follows:

$$x = Mn, \quad (10)$$

where $x \in \mathcal{R}^n$ are points in the lattice, $M \in \mathcal{R}^{n \times n}$ is the so-called lattice-generating matrix, and $n \in \mathcal{Z}^n$ is an integer vector. Since the points we are interested in belong to $\{1, \dots, N\}^n$, we require that x have integer entries. We will therefore henceforth assume that M has positive integer entries, so that $M \in \mathcal{N}^{n \times n}$. We will refer to the lattice generated by M as $\mathcal{L}(M)$.

Definition 2 (Lattice-Generated Distribution): A probability distribution over n random variables with alphabet size N each, will be called *lattice-generated*, if for some lattice $\mathcal{L}(M)$, we have

$$p_{X_{\mathcal{N}}}(x_{\mathcal{N}}) = c,$$

a constant, whenever $x_{\mathcal{N}} \in \{0, \dots, N-1\}^n \cap \mathcal{L}(M)$ and

$$p_{X_{\mathcal{N}}}(x_{\mathcal{N}}) = 0,$$

otherwise.

Now we need a few lemmas.

Lemma 1 (Bezout Identity): The following equality holds for 2-by-2 lattices.

$$\mathcal{L} \left(\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \right) = \mathcal{L} \left(\begin{bmatrix} \gcd(M_{11}, M_{12}) & 0 \\ M_{21}x + M_{22}y & \frac{M_{11}M_{22} - M_{21}M_{12}}{\gcd(M_{11}, M_{22})} \end{bmatrix} \right), \quad (11)$$

where x and y are integers found from the Bezout identity $M_{11}x + M_{12}y = \gcd(M_{11}, M_{12})$.

Proof: Follows from post-multiplication by the unimodular matrix $\begin{bmatrix} x & -M_{12}/\gcd(M_{11}, M_{12}) \\ y & M_{11}/\gcd(M_{11}, M_{12}) \end{bmatrix}$. ■

Lemma 2 (Lower Triangularization): Any lattice generating matrix with non-negative integer entries can be lower triangularized without changing the resulting lattice.

Proof: Follows from repeated use of Lemma 1 in a fashion akin to QR factorization. ■

Lemma 3 (Lattice-Generated Quasi-Uniform Distributions): A lattice-generated distribution is quasi-uniform if the lattice has a period that divides N . The latter is true if, and only if, the matrix $M^{-1}N$ has integer entries.

Proof: Assume the lattice M has a period that divides N . This is true if, and only if, for every $x \in \mathcal{L}(M)$ the point $x + Ne_i$ belongs to $\mathcal{L}(M)$ for all $i = 1, \dots, n$, where e_i is the i -th unit vector with one in the i -th position and zeros elsewhere. In other words, if there exists an integer vector n such that $Mn = x$, there should also exist an integer vector n' such that $Mn' = x + Ne_i$. But $n' = M^{-1}x + M^{-1}Ne_i = n + M^{-1}Ne_i$, which implies that $M^{-1}Ne_i$ should have integer entries for all i and establishes the second claim.

We now need to show that the resulting distribution is quasi-uniform. To this end, note that

$$p_{X_\alpha}(x_\alpha) = \sum_{x_{N-\alpha}} p_{X_N}(x_N) = c \sum_{x_{N-\alpha}, p_{X_N}(x_\alpha, x_{N-\alpha}) \neq 0} 1,$$

which implies that a distribution taking on only the values 0 and c is quasi-uniform if, and only if, for every value x_α for which $p_{X_\alpha}(x_\alpha)$ is nonzero, the number of $x_{N-\alpha}$ for which $p_{X_N}(x_\alpha, x_{N-\alpha})$ is non-zero should be constant.

Now partitioning the lattice-generating matrix according to α and $\alpha^\circ = N - \alpha$ and lower-triangularizing using Lemma 2 yields

$$\begin{bmatrix} x_\alpha \\ x_{\alpha^\circ} \end{bmatrix} = \begin{bmatrix} M_{\alpha, \alpha} & 0 \\ M_{\alpha^\circ, \alpha} & M_{\alpha^\circ, \alpha^\circ} \end{bmatrix} \begin{bmatrix} n_\alpha \\ n_{\alpha^\circ} \end{bmatrix}. \quad (12)$$

Any value of x_α that yields a non-zero $p_{X_\alpha}(x_\alpha)$ is one for which $n_\alpha = M_{\alpha, \alpha}^{-1}x_\alpha$ is an integer vector. Therefore so is the vector $M_{\alpha^\circ, \alpha}M_{\alpha, \alpha}^{-1}x_\alpha$. Thus, the number of x_{α° for which $p_{X_N}(x_\alpha, x_{\alpha^\circ})$ is nonzero is given by the number of x_{α° in $\{0, \dots, N-1\}^{|\alpha^\circ|}$ for which $M_{\alpha^\circ, \alpha^\circ}n_{\alpha^\circ} = x_{\alpha^\circ} - M_{\alpha^\circ, \alpha}M_{\alpha, \alpha}^{-1}x_\alpha$ has an integer solution in n_{α° . However, since the lattice is periodic with period dividing N , this number is independent of the integer shift $M_{\alpha^\circ, \alpha}M_{\alpha, \alpha}^{-1}x_\alpha$. ■

Lemma 3 tells us that we should focus on lattice-generating matrices such that $M^{-1}N$ has integer entries. The next lemma of this section shows us how to extract the entropies from the lattice generating matrix M .

Lemma 4 (Entropy Extraction): Consider a lattice-generated distribution with period dividing N . Then the normalized entropy of any collection of random variables X_α is given by

$$h(\alpha) = |\alpha| - \frac{\log \det M_{\alpha, \alpha}}{\log N}, \quad (13)$$

where $M_{\alpha, \alpha}$ is found from the lower triangularization of the lattice-generating matrix in (12).

Proof: The distribution is quasi-uniform and so the entropy $h(\alpha)$ is simply the log of the number of non-zero points in the distribution $p_{X_\alpha}(x_\alpha)$. The total number of points is $N^{|\alpha|}$ and the volume of the basic volume element in the lattice corresponding to the variables x_α is well known to be $\det M_{\alpha, \alpha}$. This gives the total number of non-zero points in the distribution as $N^{|\alpha|}/\det M_{\alpha, \alpha}$, which yields the desired result. ■

Remark: In a more general setting, (12) can be written as follows:

$$\begin{bmatrix} x_\alpha \\ x_{\alpha^\circ} \end{bmatrix} = \begin{bmatrix} M_\alpha \\ M_{\alpha^\circ} \end{bmatrix} \begin{bmatrix} n_\alpha \\ n_{\alpha^\circ} \end{bmatrix} \quad (14)$$

Using the Smith-McMillan form [15] for matrix M_α , (13) can be equivalently stated as:

$$h(\alpha) = |\alpha| - \frac{\log \gcd(\text{all } |\alpha| \times |\alpha| \text{ minors of } M_\alpha)}{\log N}, \quad (15)$$

which is often easier to compute.

Let Δ_n denote the space of entropy vectors obtained from lattice-generated quasi-uniform distributions.

Theorem 3 (An Inner Region for Entropic Vectors):

$\overline{\text{con}}(\Delta_n) \subseteq \Omega_n^*$ where $\overline{\text{con}}(\cdot)$ represents the convex closure.

Proof: Follows straightforwardly from the convexity of Ω_n^* . ■

Our last result states that the obtained inner region is a polytope. However, we will omit the proof for brevity.

Theorem 4: The region $\overline{\text{con}}(\Delta_n)$ is a polytope for all n .

IV. EXPLICIT CONSTRUCTIONS

Now that we have described our general construction, we will study the consequences for $n = 2, 3, 4$ random variables.

A. Two Random Variables

Without loss of generality we can assume that the 2-by-2 lattice-generating matrix is lower triangular

$$M = \begin{bmatrix} M_{11} & 0 \\ M_{21} & M_{22} \end{bmatrix}. \quad (16)$$

We should remark that for any integers M_{ij} it is always possible to find a large enough integer N and positive rational numbers γ_{ij} such that $M_{ij} = N^{\gamma_{ij}}$. Furthermore, for large enough N it follows that $\gcd(N^{\gamma_{ij}}, N^{\gamma_{kl}}) = N^{\min(\gamma_{ij}, \gamma_{kl})}$ which considerably simplifies the gcd calculation. Since we are studying normalized entropies, increasing N comes at no cost and so we will assume all of the above. We will therefore, henceforth assume that

$$M = \begin{bmatrix} N^{\gamma_{11}} & 0 \\ N^{\gamma_{21}} & N^{\gamma_{22}} \end{bmatrix}. \quad (17)$$

We first need to enforce the condition that the generated distribution be quasi-uniform. From Lemma 3 this means that the matrix

$$NM^{-1} = \begin{bmatrix} N^{1-\gamma_{11}} & 0 \\ -N^{1-\gamma_{11}-\gamma_{22}+\gamma_{21}} & N^{1-\gamma_{22}} \end{bmatrix} \quad (18)$$

must have integer entries. Since N is large enough this implies the inequalities

$$\gamma_{11} \leq 1, \quad \gamma_{22} \leq 1, \quad \gamma_{11} + \gamma_{22} \leq 1 + \gamma_{21}. \quad (19)$$

Using Lemma 4 the corresponding entropies are readily seen to be

$$\begin{aligned} h_1 &= 1 - \gamma_{11} \\ h_2 &= 1 - \min(\gamma_{21}, \gamma_{22}) \\ h_{12} &= 2 - \gamma_{11} - \gamma_{22} \end{aligned} \quad (20)$$

Thus the space $\bar{\Delta}_2$ is described by (20) along with the constraints (19).

The region $\bar{\Delta}_2$ may not be convex, due to the $\min(\cdot)$ operator in h_2 . In fact, it is not hard to show that the convex hull of (19-20) is

$$\begin{cases} h_1 = 1 - \gamma_{11}, & h_2 = 1 - \gamma_{21}, & h_{12} = 2 - \gamma_{11} - \gamma_{22} \\ 0 \leq \gamma_{11}, \gamma_{22} \leq 1, & 0 \leq \gamma_{21} \leq \gamma_{22}, & \gamma_{11} + \gamma_{22} \leq 1 + \gamma_{21} \end{cases} \quad (21)$$

Theorem 5:

$$\overline{\text{con}}\Delta_2 = \bar{\Omega}_2^*$$

Proof: Any h_{ij} obtained from (21) is an entropy vector by construction. Conversely, for any entropy vector satisfying $0 \leq h_1, h_2 \leq 1$ and $h_1, h_2 \leq h_{12} \leq h_1 + h_2$ a valid set of γ_{ij} s from (21) can be found. ■

B. Three Random Variables

Again, without loss of generality we may assume

$$M = \begin{bmatrix} N\gamma_{11} & 0 & 0 \\ N\gamma_{21} & N\gamma_{22} & 0 \\ N\gamma_{31} & N\gamma_{32} & N\gamma_{33} \end{bmatrix}. \quad (22)$$

Insisting on quasi-uniformity by forcing the elements of $M^{-1}N$ to be integers (Lemma 3) yields the linear constraints:

$$\begin{aligned} 0 &\leq \gamma_{ij} \leq 1 \\ \gamma_{ii} + \gamma_{jj} - \gamma_{ij} &\leq 1 \quad i > j \\ \gamma_{11} + \gamma_{22} + \gamma_{33} - \gamma_{21} - \gamma_{32} &\leq 1 \end{aligned} \quad (23)$$

Extracting the entropies using Lemma 4 yields

$$\begin{aligned} h_1 &= 1 - \gamma_{11} \\ h_2 &= 1 - \min(\gamma_{21}, \gamma_{22}) \\ h_3 &= 1 - \min(\gamma_{31}, \gamma_{32}, \gamma_{33}) \\ h_{12} &= 2 - \gamma_{11} - \gamma_{22} \\ h_{13} &= 2 - \gamma_{11} - \min(\gamma_{32}, \gamma_{33}) \\ h_{23} &= \begin{cases} 2 - \min(\gamma_{21} + \gamma_{32}, \gamma_{22} + \gamma_{31}, \gamma_{33} + \min(\gamma_{21}, \gamma_{22})) \\ \text{if } \gamma_{21} + \gamma_{32} \neq \gamma_{22} + \gamma_{31} \\ 2 - \gamma_{33} - \min(\gamma_{21}, \gamma_{22}) \\ \text{if } \gamma_{21} + \gamma_{32} = \gamma_{22} + \gamma_{31} \end{cases} \\ h_{123} &= 3 - \gamma_{11} - \gamma_{22} - \gamma_{33} \end{aligned} \quad (24)$$

The space $\bar{\Delta}_3$ is described by (24-23). The region is clearly non-convex. Furthermore, it is six-dimensional (since there are

six γ_{ij} parameters) whereas the entropy vectors are seven-dimensional. However, the convex hull is seven-dimensional and consists of the convex hull of four (polytope) regions. The first three regions satisfy the constraints (23), as well as,

$$\begin{aligned} h_1 &= 1 - \gamma_{11} \\ h_2 &= 1 - \gamma_{21} \\ h_3 &= 1 - \gamma_{31} \\ h_{12} &= 2 - \gamma_{11} - \gamma_{22} \\ h_{13} &= 2 - \gamma_{11} - \gamma_{32} \\ h_{123} &= 3 - \gamma_{11} - \gamma_{22} - \gamma_{33} \end{aligned} \quad (25)$$

However, region I has

$$\begin{cases} h_{23} = 2 - \gamma_{21} - \gamma_{32} \\ \gamma_{21} + \gamma_{32} < \gamma_{22} + \gamma_{31} \end{cases}, \quad (26)$$

region II has

$$\begin{cases} h_{23} = 2 - \gamma_{22} - \gamma_{31} \\ \gamma_{21} + \gamma_{32} > \gamma_{22} + \gamma_{31} \end{cases}, \quad (27)$$

and region III has

$$\begin{cases} h_{23} = 2 - \gamma_{21} - \gamma_{33} \\ \gamma_{21} + \gamma_{32} = \gamma_{22} + \gamma_{31} \end{cases}. \quad (28)$$

The fourth region is characterized by $\gamma_{21} \geq \gamma_{22}$ and $\gamma_{33} \geq \gamma_{31} \geq \gamma_{32}$. Interestingly, the convex hull of these four regions is $\bar{\Omega}_3^*$.

Theorem 6:

$$\overline{\text{con}}\Delta_3 = \bar{\Omega}_3^*$$

Proof: Clearly, $\overline{\text{con}}\Delta_3 \subseteq \bar{\Omega}_3^*$ since, by construction, all vectors in $\overline{\text{con}}\Delta_3$ are entropic.

To prove the other direction consider the region defined by

$$\begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_{12} \\ h_{23} \\ h_{31} \\ h_{123} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \\ k_6 \\ k_7 \\ k_8 \end{bmatrix}, \quad (29)$$

where $k_i \geq 0$. Each column vector in the matrix on the RHS can be seen to be generated by a lattice-generated distribution (just check!). Therefore the region must be a subset of $\bar{\Delta}_3$. If we write the above matrix equation as

$$h = \begin{bmatrix} A & a \end{bmatrix} \begin{bmatrix} k \\ k_8 \end{bmatrix} = Ak + ak_8,$$

which, since the first seven columns of the matrix on the RHS is invertible, we can further write

$$A^{-1}h - A^{-1}ak_8 = k \geq 0.$$

Computing A^{-1} and $A^{-1}ak_8$, yields

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & -1 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_{12} \\ h_{23} \\ h_{31} \\ h_{123} \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \\ k_8 \\ k_8 \\ k_8 \\ -k_8 \end{bmatrix} \quad (30)$$

The point is to show that for any entropic vector h one can find a non-negative k_8 such that above inequalities are satisfied. The first three are clearly satisfied. The next three are satisfied provided

$$k_8 \leq \min_{i,j,k} (-h_i + h_{ij} + h_{ki} - h_{ijk}), \quad (31)$$

and the last inequality if

$$k_8 \geq -\sum_i h_i + \sum_{i,j} h_{ij} - h_{ijk}. \quad (32)$$

It is straightforward to show that the upper bound on k_8 exceeds the lower bound and so the region for k_8 is non-empty. Furthermore, by the submodularity of the entropy function, the upper bound is non-negative, which implies that a non-negative k_8 can always be found. This concludes the proof. ■

C. Four Random Variables

For four random variables, we assume

$$M = \begin{bmatrix} N^{\gamma_{11}} & 0 & 0 & 0 \\ N^{\gamma_{21}} & N^{\gamma_{22}} & 0 & 0 \\ N^{\gamma_{31}} & N^{\gamma_{32}} & N^{\gamma_{33}} & 0 \\ N^{\gamma_{41}} & N^{\gamma_{42}} & N^{\gamma_{43}} & N^{\gamma_{44}} \end{bmatrix}. \quad (33)$$

The conditions from Lemma 3 translate to:

$$\begin{aligned} 0 &\leq \gamma_{ii} \leq 1, \\ \gamma_{ii} + \gamma_{jj} - \gamma_{ij} &\leq 1, \quad i > j \\ \gamma_{ii} + \gamma_{jj} + \gamma_{kk} - \gamma_{ij} - \gamma_{jk} &\leq 1, \quad i > j > k \\ \gamma_{11} + \gamma_{22} + \gamma_{33} + \gamma_{44} - \gamma_{21} - \gamma_{32} - \gamma_{43} &\leq 1, \end{aligned} \quad (34)$$

The extraction of the entropies for 4 variables is more involved as it can be seen from the following:

$$\begin{aligned} h_4 &= 1 - \min(\gamma_{41}, \gamma_{42}, \gamma_{43}, \gamma_{44}) \\ h_{14} &= 2 - \gamma_{11} - \min(\gamma_{42}, \gamma_{43}, \gamma_{44}), \\ h_{24} &= 2 - \min(\gamma_{21} + \min(\gamma_{42}, \gamma_{43}, \gamma_{44}), \\ &\quad \gamma_{22} + \min(\gamma_{41}, \gamma_{43}, \gamma_{44})) \\ h_{34} &= 2 - \min(\gamma_{31} + \min(\gamma_{42}, \gamma_{43}, \gamma_{44}), \\ &\quad \gamma_{32} + \min(\gamma_{41}, \gamma_{43}, \gamma_{44}), \\ &\quad \gamma_{33} + \min(\gamma_{41}, \gamma_{42}, \gamma_{44})) \end{aligned}$$

$$\begin{aligned} h_{124} &= 3 - \gamma_{11} - \gamma_{22} - \min(\gamma_{43}, \gamma_{44}) \\ h_{134} &= 3 - \gamma_{11} - \min(\gamma_{32} + \min(\gamma_{43}, \gamma_{44}), \\ &\quad \gamma_{33} + \min(\gamma_{42}, \gamma_{44})) \\ h_{234} &= 3 - \min(\gamma_{33} + \min(\gamma_{21} + \gamma_{42}, \gamma_{22} + \gamma_{41}), \\ &\quad \gamma_{33} + \gamma_{44} + \min(\gamma_{21}, \gamma_{22}), \\ &\quad \min(\gamma_{43}, \gamma_{44}) \\ &\quad + \min(\gamma_{21} + \gamma_{32}, \gamma_{22} + \gamma_{31})) \\ h_{1234} &= 4 - \gamma_{11} - \gamma_{22} - \gamma_{33} - \gamma_{44} \end{aligned} \quad (35)$$

These equations along with (24) give the whole set of equations for 4 variables. Note from (15) that whenever a minor is zero that case should be considered separately in the above equations. The space described by (35) is clearly non-convex and so, as in the case with three random variables, we must focus on its convex hull.

The most efficient ways of doing so and the tightness of the inner bound are under study. Nonetheless, since the region is a polytope it should allow for the computation of inner bounds on a wide class of network problems via simple linear programming.

REFERENCES

- [1] Raymond W. Yeung, "A framework for linear information inequalities," *IEEE Trans. on Information Theory*, vol. 43, no. 6, pp. 1924–1934, 1997.
- [2] Satoru Fujishige, "Polymatroidal dependence structure of a set of random variables," *Information and Control*, vol. 39, pp. 55–72, 1978.
- [3] Te Sun Han, "A uniqueness of shannon's information distance and related nonnegativity problems," *J. Comb., Inform. Syst. Sci.*, vol. 6, no. 4, pp. 320–331, 1981.
- [4] Zhen Zhang and Raymond Yeung, "On characterization of entropy function via information inequalities," *IEEE Trans. on Information Theory*, vol. 44, no. 4, pp. 1440–1452, 1998.
- [5] Zhen Zhang and Raymond Yeung, "A non-shannon-type conditional inequality of information quantities," *IEEE Trans. on Information Theory*, vol. 43, no. 6, pp. 1982–1986, 1997.
- [6] R. Dougherty, C. Freiling, and K. Zeger, "Six new non-shannon information inequalities," in *IEEE International Symposium on Information Theory (ISIT)*, 2006, pp. 233–236.
- [7] K. Makarychev, Y. Makarychev, A. Romashchenko, and N. Vereshchagin, "A new class of non-shannon-type inequalities for entropies," *Communications in Information and Systems*, vol. 2, no. 2, pp. 147–166, 2002.
- [8] Zhen Zhang, "On a new non-shannon type information inequality," *Communications in Information and Systems*, vol. 3, no. 1, pp. 47–60, 2003.
- [9] F. Matus, "Two constructions on limits of entropy functions," *IEEE Trans. on Information Theory*, vol. 53, no. 1, pp. 320–330, 2007.
- [10] R. Dougherty, C. Freiling, and K. Zeger, "Matroids, networks and non-shannon information inequalities," *IEEE Trans. on Information Theory*, submitted January 2006.
- [11] Raymond W. Yeung and Zhen Zhang, "Distributed source coding for satellite communications," *IEEE Trans. on Information Theory*, vol. 45, no. 4, pp. 1111–1120, 1999.
- [12] F. Matus and M. Studeny, "Conditional independences among four random variables I," *Combin., Prob. Comput.*, vol. 4, pp. 269–278, 1995.
- [13] Terence H. Chan, "A combinatorial approach to information inequalities," *Communications in Information and Systems*, vol. 1, no. 3, pp. 241–254, 2001.
- [14] Terence H. Chan and Raymond W. Yeung, "On a relation between information inequalities and group theory," *IEEE Trans. on Information Theory*, vol. 48, no. 7, pp. 1992–1995, 2002.
- [15] M.C. Smith, "Applications of algebraic function theory in multivariable control," *Multivariable control: New concepts and tools*, S.G. Tzafestas, Ed., Reidel Pub., 1984.